

パノラマ背景画像への移動物体画像の合成

星 野 准[†] 増 田 功[†]

多数のカメラを利用した監視においては、入力画像を重ね合わせてパノラマ画像を生成することにより、監視環境の全体的な状況把握を容易にすることが望まれている。ところが、あらかじめ作成したパノラマ画像を背景として、移動している人物画像を重ね合わせる場合のように、入力画像と背景画像の内容が異なる場合は、変換パラメータの推定が不安定になるという問題がある。本稿では、パノラマ背景画像へ移動物体画像を合成する際に、移動物体の侵入による誤差成分を抑制して、画像間の変換パラメータを安定かつ高精度に求める手法を提案する。本手法では、まず大まかなカメラの姿勢を利用して、画像間の変換パラメータの初期値を推定する。次に、重み付き最小2乗推定により、移動物体の侵入による輝度領域の誤差成分を抑制して、画像間の変換パラメータを安定に推定する。本稿では、本手法の有効性を実験により確認した結果と、画像監視への応用例を述べる。

A Robust Method for Merging Moving Object Images with a Panoramic Background

JUN'ICHI HOSHINO[†] and ISAO MASUDA[†]

One of the common method for creating panoramic images is to registrate input images using the planar perspective transformation. The planar perspective transformation can be estimated by minimizing intensity errors between a base image and an input image. However, the parameter estimation fails when the local intensity variations occur due to the moving objects. In this paper, we proposed a robust technique for merging moving object images with a panoramic background image. First, we decompose the local intensity variation into three components (i.e., translation in x and y direction, and the occlusion caused by the moving objects) by comparing the background image and the input image. Then, we apply a weighted least square method to eliminate intensity errors due to the occlusion component. The experimental results show that the accurate registration can be possible using the proposed method.

1. ま え が き

多数のカメラを利用した監視においては、カメラ台数の増加にともない、断片的な視野が増加するため、撮影環境の全体像の把握が困難になる。そのため、入力画像を重ね合わせて、パノラマ画像を生成することにより、監視環境の全体的な状況把握を強化することが望まれている¹⁾。

パノラマ画像を生成する手法としては、入力画像を平面で近似して、画像間の変換パラメータを求める手法がよく利用される^{4)~10)}。これらの手法では、画像間の輝度値の2乗誤差が最小になるように、画像間の変換パラメータを推定する。ところが、あらかじめ作成したパノラマ画像を背景として、移動している人物

画像を重ね合わせる場合には、カメラの姿勢やズーム率が異なることが多い。その場合は、入力画像と背景画像の内容が異なるために、変換パラメータの推定が不安定になるという問題がある。

このように、あらかじめ生成したパノラマ画像を背景として、移動物体が含まれる画像系列を合成する手法としては、特殊なカメラシステムを利用することも考えられる。たとえば、投影中心まわりにパン・チルト・ズームを行い、正確な回転角が獲得可能なカメラも提案されている¹⁷⁾。ところが、実際の監視環境では、すでに設置されている監視カメラを有効に活用して、簡易な変更で付加することが望まれることも多い。その場合は、画像処理部を追加するだけで利用できることが好ましい。また、通常の監視カメラでは回転角に誤差があるが、設置時に正確なキャリブレーションを行うことは運用上難しいことが多いため、画像処理の段階で誤差を吸収できることが重要となる。

[†] セコム IS 研究所画像情報処理研究室
Image Processing Department, SECOM Intelligent Systems Laboratory

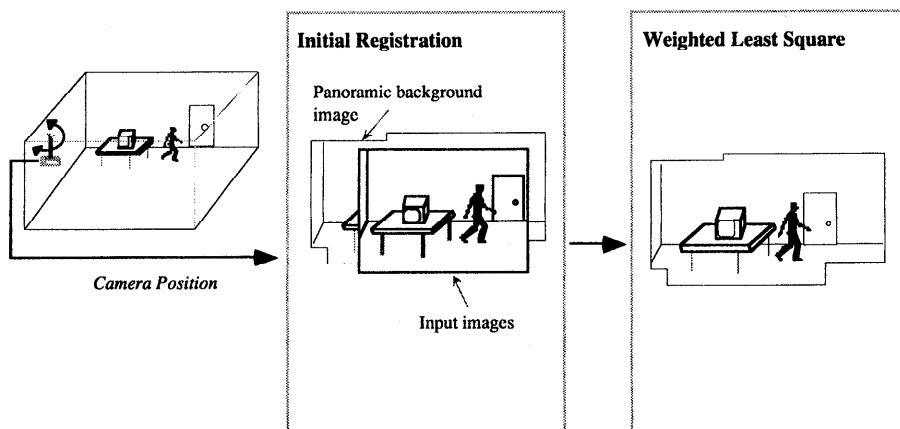


図 1 移動物体を含む画像系列からのパノラマ画像生成処理の概要
Fig. 1 Process of the panoramic image mosaicing with moving objects.

本稿では、パノラマ背景画像に移動物体を合成する際に、移動物体による誤差成分を抑制して、画像間の変換パラメータを安定かつ高精度に求める手法を提案する。本手法では、まず誤差が含まれているカメラの姿勢を利用して、画像間の変換パラメータの初期値を推定する。次に、重み付き最小 2 乗推定により、移動物体の侵入による輝度領域の誤差成分を抑制して、画像間の変換パラメータを安定かつ高精度に推定する。

本稿では、2 章でシステムの概要について述べて、3 章でカメラの回転角を利用して画像間の変換パラメータを予測する手法を示す。4 章では、移動物体領域を推定する手法と、重み付き最小 2 乗推定により変換パラメータの最適化を行う手法を示す。5 章では本手法の有効性を実験により確認した結果を示す。

2. システムの概要

2.1 射影変換によるパノラマ画像の生成

図 1 に本システムの概要を示す。本システムでは、まず適当な間隔で取得した背景画像を重ね合わせて、パノラマ背景画像を生成する。次に、パノラマ背景画像に、新しい入力画像を重ね合わせて利用者に提示する。

複数枚の画像を重ね合わせる手法としては、入力画像間の変形を射影変換で近似する方法が提案されている^{6)~13)}。このような仮定は、カメラの移動が投影中心まわりに回転する場合か、視点移動に対して背景が平面で近似できる場合に成立する。射影変換では、背景画像 $I(x_i, y_i)$ に入力画像 $I'(x'_i, y'_i)$ を重ね合わせるときに、2 次元画像面上の座標変換を次のように記述することができる。

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (1)$$

また、式 (1) を書き直すと、次のようになる。

$$\begin{aligned} x'_i &= \frac{m_0 x_i + m_1 y_i + m_2}{m_6 x_i + m_7 y_i + 1} \\ y'_i &= \frac{m_3 x_i + m_4 y_i + m_5}{m_6 x_i + m_7 y_i + 1} \end{aligned} \quad (2)$$

変換パラメータ $\{m_0 \dots m_7\}$ を自動的に求める場合は、背景画像と入力画像の輝度値の誤差が最小になるように最適化する。

$$E_m = \sum_i [I'(x'_i, y'_i) - I(x_i, y_i)]^2 = \sum_i e_i^2 \quad (3)$$

そして、変換パラメータ $\{m_0 \dots m_7\}$ を利用して、逆変換を行うことにより、入力画像 $I'(x'_i, y'_i)$ を背景画像 $I(x_i, y_i)$ に重ね合わせる。画像間の 2 乗誤差を最小化する手法としては Levenberg-Marquadt 法¹⁴⁾などの非線形パラメータ最適化手法を利用する。

2.2 本手法の概要

静止している情景から入力画像が得られた場合は、重複する領域の輝度値が一致するため、画像間の 2 乗誤差を最小化することで変換パラメータを推定することが可能である。ところが、移動物体の侵入による輝度変動が起こる場合は、正しい位置に重ね合わせても 2 乗誤差が最小にならないことが多い。そのため、非線形最適化の過程で収束が不安定になり、まったく異なる値に収束することもある。また、非線形最適化手法で安定にパラメータを推定するためには、画像間の並行移動成分などの初期値を与えることが必要となる。

ところが、画像の内容が異なる場合は、初期値を求める段階で正しい値を得ることが難しい。

本手法では、上記の問題を解決するために、次の2つのステップで画像の重ね合わせを行う。

- (1) 大まかなカメラの姿勢を利用して、画像間の変換パラメータの初期値を求める。
- (2) 画像間の輝度値の2乗誤差が最小になるように変換パラメータの最適化を行う。最適化を行う際に、あらかじめ移動物体領域を推定して、重み付き最小2乗推定を適用する。

3. 射影変換パラメータの予測

本章では、カメラの運動が投影中心まわりの回転で表される場合に、カメラの回転角から射影変換マトリクスを算出する手法を示す。

3.1 射影変換の定式化

本章では、同次座標系 (homogenous coordinates) で関係式の記述を行う¹⁵⁾。同次座標系では、 (x, y, w) に相当する点の2次元空間内の座標は $(x/w, y/w)$ となる。同様に、 (x, y, z, w) の3次元空間内の座標値は $(x/w, y/w, z/w)$ となる。

3次元空間内の点 $p = (x, y, z, w)$ と、2次元画面上の点 $u = (x, y, w)$ の関係は、3次元空間内の剛体変換 E と、透視変換 V で記述することができる。

$$u = VE p \quad (4)$$

式(4)の透視変換マトリクス V は 3×4 の行列である。

$$V = [\hat{V}0] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \quad (5)$$

また、式(4)の剛体変換 E は、カメラ間の移動を表す 3×3 の回転行列 R および、3次元の並行移動ベクトル t で表すことができる。

$$E = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \quad (6)$$

3.2 カメラ姿勢と射影変換マトリクスの関係

画像面上の点 u に対応した3次元空間内の点 p は、逆透視変換となるため、透視変換マトリクス V の逆変換で表すことができる。

$$p = \begin{bmatrix} \hat{V}^{-1}u \\ w \end{bmatrix} \quad (7)$$

ここで w は、射影的な奥行き (projective depth) であり、同次座標系において点 p が原点からどの程度

離れているかを表している。

次に、式(7)の p をカメラの原点を中心に回転させて点 p' を求める。カメラの原点を中心に回転する場合は、式(4)の剛体変換 E で $t = 0$ となるため、

$$p' = E p = \begin{bmatrix} R & 0 \\ 0^T & 1 \end{bmatrix} p = \begin{bmatrix} R \hat{V}^{-1}u \\ w \end{bmatrix} \quad (8)$$

である。よって、式(4)に示すように $u = VE p$ であるため、式(8)を代入すると、 u' と u の関係を求めることができる。

$$u' = \hat{V}' R \hat{V}^{-1} u = M_R u \quad (9)$$

式(9)の M_R はカメラが原点を中心に回転する場合の、射影変換マトリクスである。 M_R を要素に展開すると次のようになる。

$$M_R = \begin{bmatrix} r_{00} & r_{01} & f r_{02} \\ r_{10} & r_{11} & f r_{12} \\ r_{20}/f' & r_{21}/f' & f r_{22}/f' \end{bmatrix} \quad (10)$$

ただし、 r_{ij} は回転行列 R の要素である。また、 f, f' は2枚の入力画像の焦点距離であり、入力画像間で焦点距離が変わらない場合は $f = f'$ である。回転行列 R は、 x, y, z 軸を中心とした回転角 $(\theta_x, \theta_y, \theta_z)$ から、次の行列の積として求めることができる。

$$\begin{aligned} R_{\theta_x} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \\ R_{\theta_y} &= \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \\ R_{\theta_z} &= \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (11)$$

式(11)から回転行列の要素 r_{ij} を求めて、式(10)に代入することにより、射影変換マトリクス M_R を求める。

4. 射影変換パラメータの最適化

3章で推定した射影変換マトリクスには誤差が含まれる。そのため、重み付き最小2乗推定により、移動物体の侵入による誤差成分を抑制して、射影変換パラメータの最適化を行う。

4.1 処理の概要

本手法では、物体の侵入による輝度変化を抽出して、重み付き最小2乗推定を適用する。本手法は次の2つ

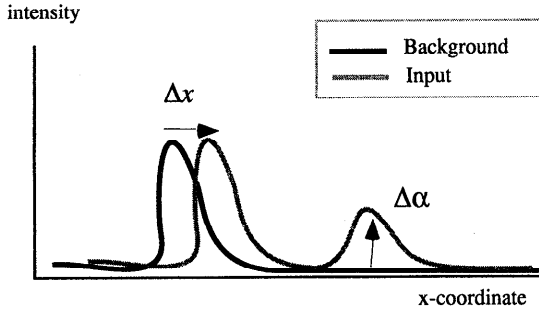


図2 背景画像の位置ずれとオクルージョンの関係

Fig. 2 Relationship between background distortion and occlusion.

のステップから構成される。

- (1) 背景画像と入力画像の輝度変化成分は、画素の位置ずれによる輝度変化と、新しい被写体の侵入による輝度変化に分類することができる。このうち、新しい被写体の侵入による局所的な輝度変化成分 $\Delta\alpha$ を推定する。
- (2) 重み付き最小2乗推定により、 $\Delta\alpha$ が大きな領域の誤差成分を抑制して、画像間のグローバルな変換パラメータを安定に推定する。

4.2 物体領域の抽出

背景画像 $I(x_i, y_i)$ と入力画像 $I'(x'_i, y'_i)$ の各画素の並行移動量を $(\Delta x, \Delta y)$ 、オクルージョンの発生による輝度変化を $\Delta\alpha$ で表す。図2では、説明を簡単にするために、 $\Delta\alpha$ と Δx の関係を示している。

次のように、各画素において2乗誤差を最小化する $(\Delta\alpha, \Delta x, \Delta y)$ を求める。

$$\sum_{i \in \Omega} W_g^2 [I(x_i - \Delta x, y_i - \Delta y) - I'(x'_i, y'_i) - \Delta\alpha]^2 \quad (12)$$

W_g はガウシアンウィンドウである。 Ω は $(\Delta\alpha, \Delta x, \Delta y)$ を求める点を中心とした近傍領域であり、ガウシアンウィンドウ W_g の領域と一致する。また、 $(\Delta\alpha, \Delta x, \Delta y)$ の変化は Ω の内部で局所的に滑らかであると仮定する。式(12)より $\Delta\alpha$ を求めると次のようになる(付録)。

$$\begin{aligned} \Delta\alpha = & \left(- \sum I_t \sum I_x I_y \sum I_x I_y \right. \\ & + \sum I_x I_t \sum I_x I_y \sum I_y \\ & - \sum I_x \sum I_x I_y \sum I_y I_t \\ & \sum I_x^2 \sum I_y \sum I_y I_t \\ & - \sum I_x \sum I_x I_t \sum I_y^2 \\ & \left. + \sum I_t \sum I_x^2 \sum I_y^2 \right) / P \end{aligned} \quad (13)$$

$$\begin{aligned} P = & - \sum I_x I_y \sum I_x I_y \\ & + 2 \sum I_x I_y \sum I_x \sum I_y \\ & - \sum I_x^2 \sum I_y \sum I_y \\ & + \sum I_y^2 \sum I_x \sum I_x + \sum I_x^2 \sum I_y^2 \end{aligned}$$

ここで、近傍領域 Ω はガウシアンウィンドウ W_g の大きさと一致するため、移動量が近傍領域 Ω より大きい場合には、異なる値が算出されるという問題がある。そのため、本稿ではピラミッド画像を利用したマルチスケールの推定を行う^{4),5)}。本手法では、解像度の低い画像で大まかな値を推定して、解像度の高い画像で精度の高い値の推定を行うことで、逐次的にパラメータの推定を更新する。原画像では移動量が近傍領域 Ω より大きい場合でも、解像度の低い画像においては見かけ上の移動量が小さくなるため、正しい値を求めることが可能となる。

図3にマルチスケール処理のブロック図を示す。図3の w はガウシアンフィルタ、 $\downarrow 2$ はダウンサンプリングを表す。本稿では、ピラミッド生成と拡大処理で、ガウシアンフィルタを利用した。2次元画像にガウシアンフィルタを適用する場合は、 x および y 方向に分離して計算することができる。そのため、ガウシアンフィルタの係数は $w = [0.0625, 0.25, 0.375, 0.25, 0.0625]$ を利用した。レベル l における画像の大きさは $1/2^l$ である。

本手法では、次のようなステップで処理を行う。

- (1) まず、最も解像度の低い画像で $(\Delta x_l, \Delta y_l, \Delta\alpha_l)$ を推定する。図3では④の処理に相当する。そして、次の階層に拡大処理を行い、 $(\Delta x_{l-1}, \Delta y_{l-1}, \Delta\alpha_{l-1})$ を求める。

$$\Delta x_{l-1} = 2EXPAND(\Delta x_l)$$

$$\Delta y_{l-1} = 2EXPAND(\Delta y_l)$$

$$\Delta\alpha_{l-1} = 2EXPAND(\Delta\alpha_l)$$

EXPAND関数はレベル i からレベル $(i-1)$ へ拡大処理を行うことを示している。本稿では、ガウシアンフィルタを用いて拡大処理を行っている。

- (2) 次に、 $(\Delta x_{l-1}, \Delta y_{l-1}, \Delta\alpha_{l-1})$ を利用して、残差成分 $(\Delta x_{l-1}^{res}, \Delta y_{l-1}^{res}, \Delta\alpha_{l-1}^{res})$ を求める。残差成分 $(\Delta x_{l-1}^{res}, \Delta y_{l-1}^{res}, \Delta\alpha_{l-1}^{res})$ の推定は、次のようなステップで行う。

- (a) 入力画像 $I(x, y)$ にワーブ処理⑤を適用する。ワーブ処理⑤は、入力画像 $I(x, y)$ の画素を推定値 $(\Delta x_{l-1}, \Delta y_{l-1}, \Delta\alpha_{l-1})$ だけ移動する処理である。ワーブ処理⑤

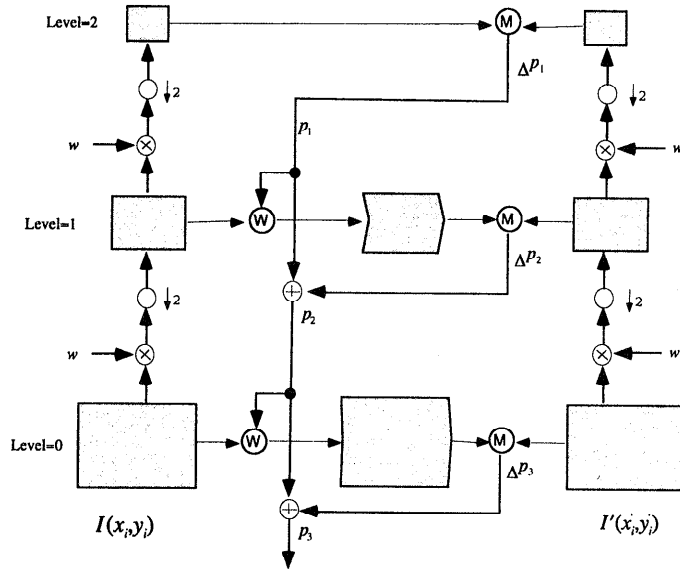


図 3 多重解像度のパラメータ推定手法のブロック図

Fig. 3 Process of the coarse-to-fine parameter estimation.

の過程でサブピクセル位置の輝度値を求めるために、双3次補間 (bicubic interpolation) を利用した¹⁴⁾。

- (b) ワープ処理Ⓢを行った入力画像 $I(x, y)$ と入力画像 $I'(x', y')$ から、残差成分 $(\Delta x_{i-1}^{res}, \Delta y_{i-1}^{res}, \Delta \alpha_{i-1}^{res})$ を求める。

- (3) 推定値 $(\Delta x_{i-1}, \Delta y_{i-1}, \Delta \alpha_{i-1})$ に残差成分 $(\Delta x_{i-1}^{res}, \Delta y_{i-1}^{res}, \Delta \alpha_{i-1}^{res})$ を加えて推定値を更新する。

$$\Delta x'_{i-1} = \Delta x_{i-1} + \Delta x_{i-1}^{res}$$

$$\Delta y'_{i-1} = \Delta y_{i-1} + \Delta y_{i-1}^{res}$$

$$\Delta \alpha'_{i-1} = \Delta \alpha_{i-1} + \Delta \alpha_{i-1}^{res}$$

画像間の移動成分を求めるときに、オクルージョンの影響を抑制する手法としては、ロバスト推定法を利用した手法が提案されている¹⁶⁾。ところが、ロバスト推定法を適用する場合は、運動パラメータと重み係数が掛け合わされた非線形の誤差関数を、数値計算法により同時に最適化を行うことが必要である。また、輝度値が変化する画素の割合が50%以下であることも必要となる。本手法では、輝度の変化量 $\Delta \alpha$ を明示的にパラメータ化するため、解析的に推定値を求めることができる。また、視点位置の変化や、照明条件の変化により、画像面の全体で明るさが変化する場合でも、安定に推定値を求めることができる。

4.3 重み付き最小2乗推定

次に、 $\Delta \alpha$ を利用して射影変換パラメータを安定に求める手法を示す。 $\Delta \alpha$ が大きな箇所では新しい被写

体が侵入するなどの変化が起こっていると仮定することができる。本稿では、移動物体領域の誤差の影響を軽減する方法として、重み付き最小2乗推定 (weighted least square)¹⁴⁾ を利用する。

$$E_m = \sum_i w(x'_i, y'_i) [I'(x'_i, y'_i) - I(x_i, y_i)]^2$$

$$= \sum_i e_i^2 \quad (14)$$

式 (14) の重み関数 $w(x'_i, y'_i)$ の性質は、 $\Delta \alpha$ が大きな領域では $w(x'_i, y'_i) \approx 0$ であり、 $\Delta \alpha$ が小さな領域では $w(x'_i, y'_i) \approx 1$ となるのが望ましい。本稿では、これらの性質を持つ典型的な関数として、次の重み関数を用いる。

$$w(x'_i, y'_i) = \begin{cases} (1 - (|\Delta \alpha|/\sigma)^2)^2 & \text{if } |\Delta \alpha| < \sigma \\ 0 & \text{if } |\Delta \alpha| > \sigma \end{cases} \quad (15)$$

σ は移動物体が存在しない場合の $|\Delta \alpha_i|$ の平均を μ_α 、標準偏差を σ_α とするときに $\sigma = \mu_\alpha + 3\sigma_\alpha$ とする。

式 (15) において、重み係数を1あるいは0に設定した場合は、 $\Delta \alpha$ に基づいてしきい値処理を行う手法に相当する。ところが、実験では背景画像と入力画像を取得する際の時間経過により、全体的に明るさが変動する場合もあった。そのため、全体的な明るさの変動の影響が少なく、移動物体領域の変動を抑制するという点で、連続的に重み係数が変化の方が安定である。

$I'(x'_i, y'_i)$ と $I(x_i, y_i)$ の2乗誤差を最小化する

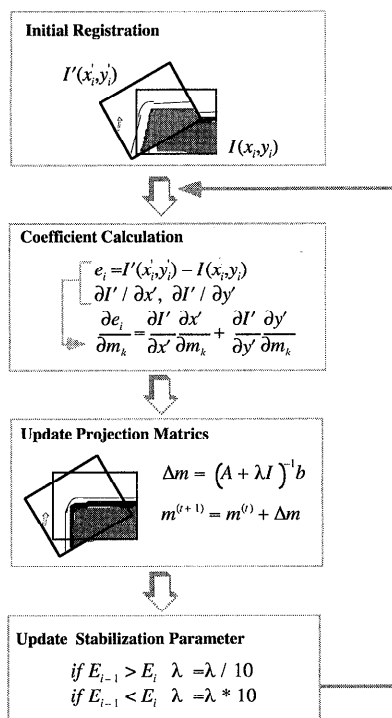


図4 射影変換マトリクスの更新の過程

Fig. 4 Update process of the projection matrix.

手法として、非線形パラメータ最適化手法である Levenberg-Marquadt 法¹⁴⁾ を利用する。Levenberg-Marquadt 法は、最急降下 (gradient descent) 法と比べて、少ない繰返し回数で収束することが知られている。Levenberg-Marquadt 法を適用するためには、 $\{m_0 \dots m_7\}$ に関する e^2 の偏微分を求めることが必要となる。

$$\frac{\partial e_i}{\partial m_0} = w(x'_i, y'_i)^{1/2} \frac{x_i}{D_i} \frac{\partial I'}{\partial x'}, \dots,$$

$$\frac{\partial e_i}{\partial m_7} = -w(x'_i, y'_i)^{1/2} \frac{y_i}{D_i} \left(x'_i \frac{\partial I'}{\partial x'} + y'_i \frac{\partial I'}{\partial y'} \right) \quad (16)$$

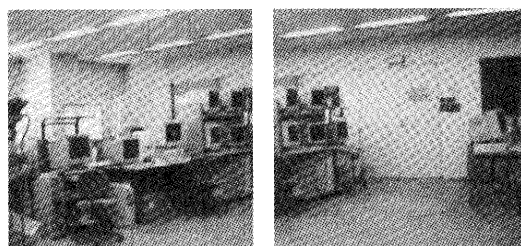
ただし、 $D_i = m_6 x_i - m_7 y_i + 1$ であり、 $(\partial I' / \partial x', \partial I' / \partial y')$ は (x'_i, y'_i) における I' の微分値である。これらの偏微分係数から、Hessian マトリックス A と、グラディエントベクトル b を求める。

$$a_{kl} = \sum_i \frac{\partial e_i}{\partial m_k} \frac{\partial e_i}{\partial m_l}, b_k = -2 \sum_i e_i \frac{\partial e_i}{\partial m_k} \quad (17)$$

そして $m = [m_0 \dots m_7]$ の推定値を次の量だけ更新する。

$$\Delta m = (A + \lambda I)^{-1} b \quad (18)$$

$$m^{t+1} = m^{(t)} + \Delta m \quad (19)$$

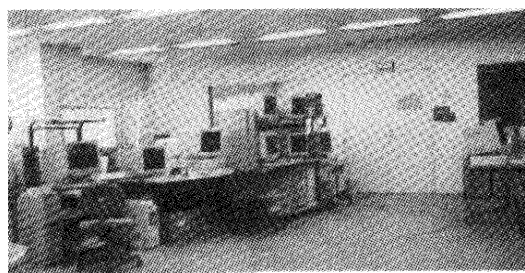


(a) 入力画像 1

(a) Input image 1

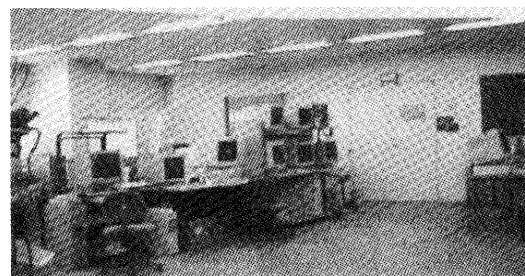
(b) 入力画像 2

(b) Input image 2



(c) 変換パラメータの予測による重ね合せ

(c) Registration using the camera rotation



(d) 変換パラメータの最適化

(d) After optimization

図5 パノラマ画像生成結果

Fig. 5 Experimental result of panoramic image mosaicing.

ここで λ は安定化係数であり、繰返しの過程で動的に更新される。また (t) は t 回目の繰返しであることを示す。図4に示すように、次のステップで処理を行う。

- (1) 背景画像 $I(x_i, y_i)$ の各画素に対応する、入力画像 $I'(x'_i, y'_i)$ の画素 (x'_i, y'_i) を求める。次に、対応する画素間の誤差 e_i と、微分値 $(\partial I' / \partial x', \partial I' / \partial y')$ を求めて、式 (16) および式 (17) を計算する。また λ の初期値を設定する。
- (2) 式 (18), (19) より射影変換マトリックス $m(t)$ の値を更新する。
- (3) 誤差 $E^{(t)}$ の値が $E^{(t-1)}$ より小さければ、 λ の値を減少させる。また、誤差 $E^{(t)}$ の値が $E^{(t-1)}$ より大きければ、 λ の値を増加させる。
- (4) 誤差の変化分 $|E^{(t)} - E^{(t-1)}|$ の値がしきい値

表 1 図 5 の射影変換パラメータ
Table 1 Projection parameters of Fig. 5.

	m_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
predicted coefficients	1.000	0.000	-160.77	0.000	1.035	0.000	0.00045	0.000	1.000
updated coefficients	1.096	-0.004	-166.01	0.055	1.0667	-13.11	0.0031	-0.00002	1.000

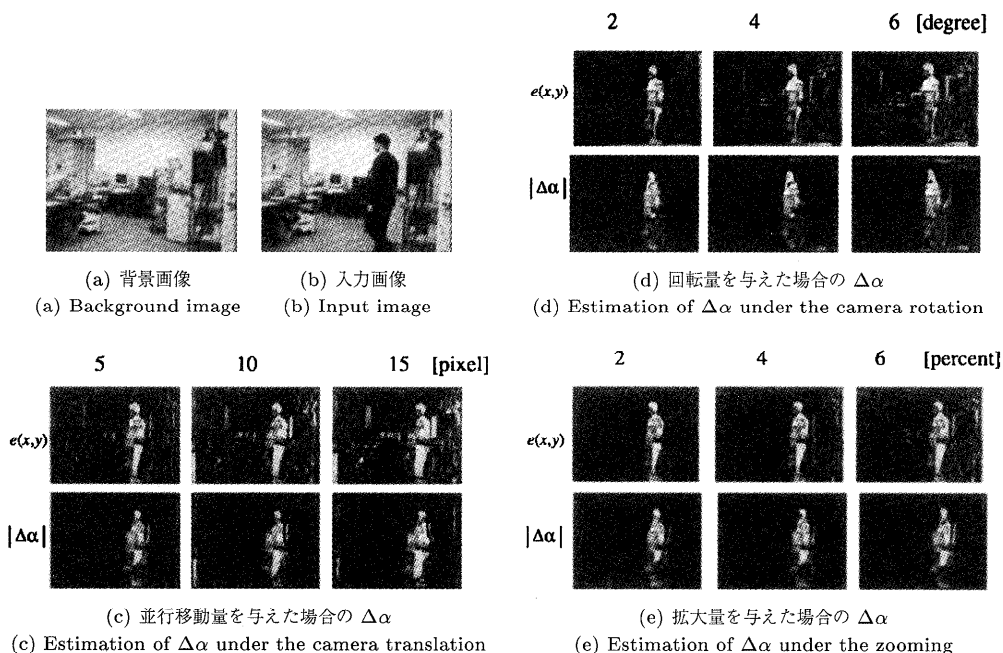


図 6 背景画像の位置ずれ量と物体抽出の安定性
Fig. 6 Stability of the object detection using $\Delta\alpha$.

よりも小さくなった場合か、決められた回数を
超えた場合に終了する。

本稿では λ の初期値を 0.001 として、減少させる
場合は 1/10, 増加する場合は 10 倍にした。終了条件
を $|E^{(t)} - E^{(t-1)}| < 0.1$ とした。また、500 回以上繰
り返しても収束しない場合は終了とした。

以上の手法により、移動物体の侵入による誤差成分
を抑制して、変換パラメータの最適化を行う。次に、
実験により本手法の有効性を確認した結果を示す。

5. 実験結果

5.1 射影変換パラメータの初期値の推定

誤差のあるカメラの回転角を利用して射影変換パラ
メータの初期値が推定できることを確認するために、
移動物体が存在しない入力画像を用いて実験を行った。
実験では、計算機からコマンド入力により角度やズー
ムを設定できる市販のカメラを利用した。通常の監視

カメラでも利用できることを想定しているため、カメ
ラの正確な角度を読み取るためのエンコーダなどは利
用していない。また、投影中心とカメラの機械的な回
転中心は、正確には一致していない。

図 5 の (a), (b) は投影中心を通る垂直軸 (y 軸)
を中心に回転させた場合の入力画像である。図 5 (c) は
 y 軸まわりの回転角を 15 度として、変換パラメータ
の初期値を推定して重ね合わせたところである。回転
角の誤差のために、つなぎ目の部分でずれて見えるこ
とが分かる。図 5 (d) は (c) で求めた変換パラメータ
を初期値として、2.1 節で述べたパラメータの最適化
を行った結果である。(d) ではつなぎ目の部分の歪み
が低減していることが分かる。表 1 に予測した変換
パラメータと、最適化を行った後の変換パラメータを
示す。

5.2 $\Delta\alpha$ の安定性の確認

$\Delta\alpha$ 推定の安定性を確認するために、既知の位置ずれ



(a) 背景画像系列
(a) Background image sequence



(b) (a) の背景画像系列から生成したパノラマ背景画像
(b) Panoramic background image

図 7 パノラマ背景画像の生成

Fig. 7 Creating panoramic background image.

を与えて実験を行った。入力画像の大きさは 320×240 画素である。また、 $\Delta\alpha$ と比較するために、それぞれの位置ずれ量に対して、画像間の輝度の差分画像を示した。図 6 (a) は基準画像、(b) は入力画像である。(c) は、基準画像と入力画像の相対的な位置を、1~15 ピクセル変化させて、 α の値を表示したところである。また、(d) は相対的な角度を 2~6 度回転させた場合、(e) は 2~6% の拡大量を与えたところである。画像間の差分を求めた場合には、画像間の位置ずれ量が増加した場合に、画面全体に差分が現れることが分かる。それに対して、 $\Delta\alpha$ では被写体領域が安定に推定できている。

5.3 移動物体画像の合成

図 7 は 3 枚の入力画像から生成されたパノラマ背景画像を生成したところである。(a) は背景画像系列、(b) は (a) から生成されたパノラマ画像である。図 8 (a) は移動する人物の拡大画像である。(a) では一番左側の画像を基準に 10 度カメラを回転させて、2 倍の拡大画像を取得した。図 8 (b) は $\Delta\alpha$ を求めたところである。(c) はカメラの回転角を利用して重ね合わせたところである。回転角の誤差のために、位置がずれていることが分かる。(d) は重み付き最小 2 乗法を利用して、(a) の入力画像を重ね合わせたところである。入力画像の位置ずれが軽減されていることが分かる。

6. む す び

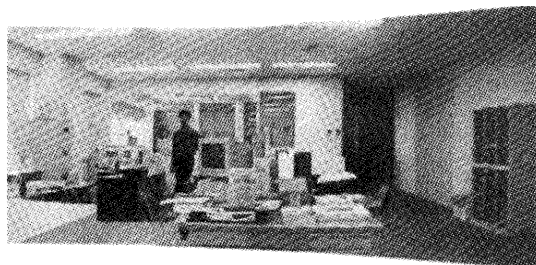
本稿では、パノラマ背景画像に移動物体を合成する



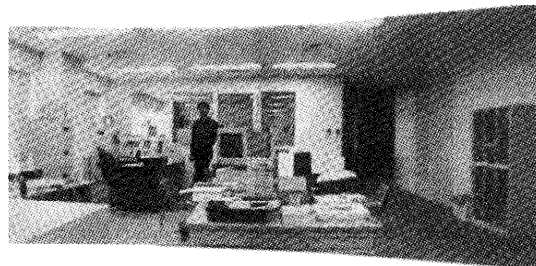
(a) 拡大画像
(a) Input zoom image



(b) $\Delta\alpha$ の推定結果
(b) Estimation of $\Delta\alpha$



(c) カメラの回転角を利用して重ね合わせた結果
(c) Merging a moving object image using the camera rotation



(d) 重み付き最小 2 乗法を適用した結果
(d) Merging a moving object image onto a panoramic background

図 8 背景画像と入力画像が異なる場合の重ね合せ

Fig. 8 Registration of the panoramic background and an input image.

際に、移動物体による誤差成分を抑制して、画像間の変換パラメータを安定かつ高精度に求める手法を提案した。本手法では、まず大まかなカメラの姿勢を利用して、画像間の変換パラメータの初期値を推定する。次に、重み付き最小 2 乗推定により、移動物体の侵入による輝度領域の誤差成分を抑制して、画像間の変換パラメータを安定に推定する。本稿では、実験により、画像間の位置ずれ量が増加した場合でも、安定に被写体領域が推定できることを確認した。また、重み付き最小 2 乗推定により、画像間の変換パラメータが安定に推定できることを確認した。

今後の研究課題としては、全体的な視野を生成するのみでなく、特定の被写体をクローズアップしたり、

異なる方向から見た視野を生成する機能を実現したりすることが考えられる。

参 考 文 献

- 1) 星野准一, 増田 功: 予測・観測画像間の誤差最小化によるパノラマ動画像の生成, 信学論, Vol.J81-D-II, No.11, pp.2565-2573 (1998).
- 2) 星野准一, 増田 功: 監視画像からの複数移動物体の抽出, 画像電子学会年次大会, pp.38-39 (1996).
- 3) 増田 功: 異常検出における画像処理技術への期待, 画像電子学会年次大会予稿, pp.45-48 (1994).
- 4) Bergen, J.R., et al.: Computing two motions from three frames, Technical Report, David Sarnoff Research Center (Apr. 1990).
- 5) Quam, L.H.: Hierarchical warp stereo, *Leadings in Computer Vision*, pp.80-66, Los Altos, CA, Morgan Kaufmann (1987).
- 6) Brown, L.G.: A Survey of image registration techniques, *ACM Comput. Surv.*, Vol.24, No.4, pp.325-376 (1992).
- 7) Perlant, F.P. and Mckeown, D.M.: Scene registration in aerial image analysis survey, *Photogrammetric Eng. Remote Sensing*, Vol.56, No.4, pp.481-493 (Apr. 1990).
- 8) Szeliski, R.: Video mosaics for virtual environments, *IEEE Computer Graphics and Applications*, Vol.16, No.2, pp.20-30 (1996).
- 9) Szeliski, R. and Shum, H.Y.: Creating full view panoramic image mosaic and environment Maps, *Proc. SIGGRAPH '97*, pp.251-258 (1997).
- 10) Mann, S. and Picard, R.: Video orbits of the projective group: A new perspective on image mosaicing, MIT Media Lab Vision and Modeling Group Technical Report #338 (1995).
- 11) Hartley, R.I.: Self-calibration from multiple views of a rotating camera, *3rd European Conf. on Computer Vision (ECCV '94)*, Vol.1, pp.471-478 (May 1994).
- 12) Irani, M., Anandan, P. and Hsu, S.: Mosaic based representation of video sequences and their applications, *5th Int. Conf. on Computer Vision (ICCV '95)*, pp.605-611 (June 1995).
- 13) Szeliski, R. and Coughlan, J.: Spline-based image registration, DEC Technical Report, CRL 94/1 (Apr. 1994).
- 14) Press, W.H., et al.: *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition, Cambridge University Press, Cambridge, England (1992).
- 15) Foley, J., et al.: *Computer Graphics: Principles and Practice*, 2nd edition, Addison-Wesley, Reading, MA (1990).
- 16) Sawney, H.S. and Ayer, S.: Compact representation of videos through dominant and multiple motion estimation, *IEEE PAMI*, Vol.18, No.8, pp.814-830 (1996).
- 17) 和田俊和, 浮田宗伯, 松山隆司: 固定型パンチルトズームカメラとその応用, 信学論, Vol.J81-D-II, No.6, pp.1182-1193 (1998).

付 録

A.1 $\Delta\alpha$ の導出

式 (12) から $\Delta\alpha$ を求める. 式 (12) をテーラー展開すると,

$$E = \sum_{i \in \Omega} W_g^2 \left[I(x_i, y_i) - I'(x'_i, y'_i) - \Delta x \frac{\partial}{\partial x} I(x_i, y_i) - \Delta y \frac{\partial}{\partial y} I(x_i, y_i) - \Delta\alpha \right]^2 \quad (20)$$

である. 式 (21) において $\partial I(x_i, y_i)/\partial x$ および $\partial I(x_i, y_i)/\partial y$ を I_x , I_y と置く. また $I_t = I(x_i, y_i) - I'(x'_i, y'_i)$ と定義すると, 次のように書くことができる.

$$E = \sum_{i \in \Omega} W_g^2 [I_t - \Delta x I_x - \Delta y I_y - \Delta\alpha]^2 \quad (21)$$

そして, E が最小になる $(\Delta\alpha, \Delta x, \Delta y)$ を求める:

$$\begin{aligned} \sum_{i \in \Omega} [I_t - \Delta x I_x - \Delta y I_y - \Delta\alpha] I_x &= 0 \\ \sum_{i \in \Omega} [I_t - \Delta x I_x - \Delta y I_y - \Delta\alpha] I_y &= 0 \\ \sum_{i \in \Omega} [I_t - \Delta x I_x - \Delta y I_y - \Delta\alpha] &= 0 \end{aligned} \quad (22)$$

式 (23) をマトリックス形式に書き直すと次のようになる:

$$\begin{bmatrix} \sum I_x^2 & \sum I_x I_y & \sum I_x \\ \sum I_x I_y & \sum I_y^2 & \sum I_y \\ \sum I_x & \sum I_y & \sum 1 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\alpha \end{bmatrix} = \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \\ \sum I_t \end{bmatrix} \quad (23)$$

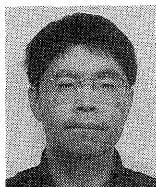
よって,

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\alpha \end{bmatrix} = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y & \sum I_x \\ \sum I_x I_y & \sum I_y^2 & \sum I_y \\ \sum I_x & \sum I_y & \sum 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \\ \sum I_t \end{bmatrix} \quad (24)$$

である. $[\]^{-1}$ は逆行列を表す. 式 (24) を展開すると式 (13) の $\Delta\alpha$ が得られる.

(平成 10 年 9 月 5 日受付)

(平成 11 年 2 月 16 日採録)



星野 准一 (正会員)

平成元年早稲田大学理工学部電子通信学科卒業. 同年, セコム IS 研究所入所. 画像情報処理研究室にて, 室内 3 次元復元, 運動解析, 人物検出, 画像生成の研究開発に従事. 平成 3 年から平成 5 年 MIT メディアラボ客員研究員. IEEE, 電子情報通信学会, 画像電子学会各会員.



増田 功 (正会員)

昭和 39 年九州大学工学部電子工学科卒業. 昭和 41 年同大学院修士課程修了. 日本電信電話公社 (現 NTT) 電気通信研究所に入所. 以来, 音声情報処理, 文字認識, 画像処理の研究に従事. 工学博士. 昭和 62 年 10 月よりセコム (株) IS 研究所. IEEE, 電子情報通信学会, 画像電子学会, 映像情報メディア学会, 人工知能学会各会員.